PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07H 17/00, C12N 15/00, 5/00

A1

(11) International Publication Number: WO 98/20016

(43) International Publication Date: 14 May 1998 (14.05.98)

(21) International Application Number: PCT/US97/20016

(22) International Filing Date: 31 October 1997 (31.10.97)

(30) Priority Data:

60/030,279	4 November 1996 (04.11.96)	US
60/049,018	9 June 1997 (09.06.97)	US

(71) Applicant *(for all designated States except US)*: SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: ESSER, Klaus, Max [US/US]; 200 Hickey Drive, Downingtown, PA 19335 (US). CHAN, John, Y. [US/US]; Apartment 210, 570 West Dekalb Pike, King of Prussia, PA 19406 (US). DABROWSKI-AMARAL, Christine, Ellen [US/US]; 4090 Moscow Road, Parkesburg, PA 19365 (US). DELVECCHIO, Alfred, Michael [US/US]; 668 Metro Court, West Chester, PA 19380 (US). DILLON, Susan, B. [US/US]; 1209 Tullamore Circle, Chester Springs, PA 19425 (US). LEARY, Jeffrey, Joseph [-/US]; 521 General Scott Road, Wayne, PA 19087 (US).

(74) Agents: GEIGER, Kathleen, W. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NOVEL CODING SEQUENCES FROM HERPES SIMPLEX VIRUS TYPE-2

(57) Abstract

This invention relates to newly identified HSV-2 polynucleotides, polypeptides encoded by such polynucleotides, the uses of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides and recombinant host cells transformed with the polynucleotides. This invention also relates to inhibiting or activating the biosynthesis or action of such polynucleotides or polypeptides and the use of such inhibitors or activators in therapy.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NOVEL CODING SEQUENCES FROM HERPES SIMPLEX VIRUS TYPE-2

Field of the Invention:

This invention relates to newly identified Herpes Simplex Virus type 2 (HSV-2) polynucleotides, polypeptides encoded by such polynucleotides, the uses of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides and recombinant host cells transformed with the polynucleotides. This invention also relates to inhibiting the biosynthesis or action of such polynucleotides or polypeptides and to the use of such inhibitors in therapy of viral infections or related diseases.

Background of the Invention:

The herpes viruses consist of large icosahedral enveloped virions containing linear double stranded DNA genomes. Currently, eight human herpes viruses have been isolated and are known to be responsible for a variety of disease states, from sub-clinical infections to fatal disease in the immuno compromised. One human herpes virus, herpes simplex virus type 2, designated HSV-2, is usually acquired through sexual contact giving rise to the condition known as genital herpes. The frequency of recurrence of secondary genital herpes ranges between one and six times per year per infected individual. It is estimated that genital HSV-2 infections occur in ten to sixty million individuals in the USA. Less frequently, HSV-2 infection results in herpes labialis, seen as cold sores.

General information about HSV-2 may be found in various treatises such as, Herpes Simplex Viruses, In: "Field's Virology", 3rd ed., Lippincott-Raven Publ, pp2297-2342 (1996); Magder, L.S., et al., New. England J. Med. 321:7-12 (1989); and "The Human Herpes viruses", Roizman, B. et al., eds. Raven Press, New York, (1993), the contents of which are incorporated herein by reference for purposes of background.

Currently, there are no vaccines available to protect against HSV-2 infection. Individuals continue to become infected by the virus and no completely satisfactory anti-viral agents or vaccines are available. Thus HSV-2 presents a major public health problem. There is a need for prophylactic and therapeutic vaccines as well as a method of identifying anti-HSV-2 agents and for reagents useful in such methods. There is a need for a method of identifying compounds which modulate the activity of HSV-2 polynucleotides and proteins and which affect the ability of the virus to replicate and produce multiple infectious virions in an infected cell. There is a need for methods of, and kits for, distinguishing HSV-2 infections from other herpes virus infections.

Brief Description of the Invention:

Toward these ends, it is an object of the present invention to provide polypeptides, *inter alia*, that have been identified as novel HSV-2 polypeptides by comparison between the amino acid sequences set out in Tables 1-4 and known amino acid sequences of proteins of other viruses such as herpes simplex virus type-1 (HSV-1).

It is a further object of the invention, to provide polynucleotides that encode HSV-2 proteins, particularly polynucleotides that encode the polypeptides encoded by the Open Reading Frames (ORFs) provided herein, or fragments, analogs or derivatives thereof.

In a particularly preferred embodiment of this aspect of the invention the polynucleotides comprise any of the regions encoding HSV-2 proteins in the sequences set out in Tables 1-4, including fragments, analogs or derivatives thereof.

In another particularly preferred embodiment of the present invention, there is a novel HSV-2 protein comprising any of the amino acid sequences shown in Table 1, or fragments, analogues or derivatives thereof.

In accordance with the invention there is provided an isolated nucleic acid molecule encoding a polypeptide expressible by the *HSV-2* polynucleotide contained in the deposited HSV-2 strain, SB5.

In accordance with the invention there are provided isolated nucleic acid molecules encoding HSV-2 proteins, nucleic acid molecules such as, mRNAs, cDNAs, genomic DNAs and, in further embodiments of this aspect of the invention, biologically, diagnostically, clinically or therapeutically useful variants, analogs or derivatives thereof, or fragments thereof, including fragments of the variants, analogs and derivatives.

Among the particularly preferred embodiments of this aspect of the invention are naturally occurring allelic variants of HSV-2 proteins.

In accordance with this aspect of the invention there are provided novel polypeptides of HSV-2 origin as well as biologically, diagnostically or therapeutically useful fragments thereof, as well as variants, derivatives and analogs of the foregoing and fragments thereof.

In accordance with certain preferred embodiments of this and other aspects of the invention there are probes that hybridize to HSV-2 sequences useful for detection of viral infection.

It also is an object of the invention to provide HSV-2 polypeptides or fragments thereof that may be employed for therapeutic or prophylactic purposes, for example, to treat disease, including treatment by conferring host immunity against viral infections, or as an antiviral agent or a vaccine.

In accordance with another aspect of the present invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization.

Among the particularly preferred embodiments of this aspect of the invention are variants of HSV-2 polypeptides encoded by naturally occurring alleles of HSV-2 genes for therapeutic or prophylactic use.

It is another object of the invention to provide a process for producing the aforementioned polypeptides, polypeptide fragments, variants and derivatives, fragments of the variants and derivatives, and analogs thereof.

In a preferred embodiment of this aspect of the invention there are provided methods for producing the aforementioned HSV-2 polypeptides comprising culturing host cells having expressibly incorporated therein an exogenously-derived HSV-2 encoding polynucleotide under conditions for expression of HSV-2 in the host and then recovering the expressed polypeptide.

In accordance with another object of the invention there are provided products, compositions, processes and methods that utilize the aforementioned polypeptides and polynucleotides, *inter alia,* for research, biological, clinical, diagnostic, prophylatic and therapeutic purposes.

In accordance with yet another aspect of the present invention, there are provided inhibitors of such polypeptides, useful as antiviral agents. In particular, there are provided antibodies against such polypeptides. In certain particularly preferred embodiments in this regard, the antibodies are selective for HSV-2.

In a further aspect of the invention there are provided compositions comprising a HSV-2 polynucleotide or HSV-2 polypeptide for administration to cells *in vitro*, to cells *ex vivo* and to cells *in vivo*, or to a multicellular organism. In certain preferred embodiments of this aspect of the invention, the compositions comprise a HSV-2 polynucleotide for expression of a HSV-2 polypeptide in a host organism to raise an immunological response, preferably to raise immunity in such host against HSV-2 or related organisms.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

Detailed Description of the Invention:

Tables 1-3 show the nucleotide sequences of one strand of "contigs," prepared by assembling sequences derived by sequencing HSV-2, Strain SB5, DNA. Collectively, the contigs herein represent between 85% to over 90% of the genome of this organism. Each of Table 1, 2 and 3 represents a separate sequencing of the HSV-2, SB5, DNA.

Tables 1-3 also show the nucleotide sequences of open reading frames (ORFs), which are deduced DNA coding sequences present within each contig. Tables 1-4 also show the deduced amino acid sequences of polypeptides encoded by these ORFs and sequence homologies to proteins in the NCBI non-redundant protein database.

Each ORF represents a HSV-2 gene although in some cases, a given ORF may actually have been derived from a gene that is longer than the ORF.

Each of the DNA sequences provided herein may be used in the discovery and development of antiviral compounds. For sequences containing an open reading frame (ORF) with appropriate initiation and termination codons, the encoded protein upon expression can be used as a target for the screening of antiviral drugs. Additionally, the DNA sequences encoding preferably the amino terminal regions of the encoded protein, or regions immediately upstream therefrom, can be used to construct antisense sequences to control the expression of the coding sequence of interest. Furthermore, many of the sequences disclosed herein also provide regions upstream and downstream from the encoding sequence. These sequences are useful as a source of regulatory elements for the control of viral gene expression. Such sequences are conveniently isolated by restriction enzyme action or synthesized chemically and introduced, for example, into promoter identification strains. These strains contain a reporter structural gene sequence located downstream from a restriction site such that if an active promoter is inserted, the reporter gene will be expressed.

Although each of the sequences may be employed as described above, this invention also provides several means for identifying particularly useful target genes. The first of these approaches entails searching appropriate databases for sequence matches in related organisms. Thus, if a homologue exists, the HSV-2-like form of this gene would likely play an analogous role. For example, a HSV-2 protein identified as homologous to a cell surface protein in another organism would be useful as a vaccine candidate. To the extent such homologies have been identified for the sequences disclosed herein they are reported along with the encoding sequence in the Tables.

A number of methods can be used to identify genes which are essential to survival *per se*, or essential to the establishment/maintenance of an infection. Identification of an

ORF unknown by one of these methods yields additional information about its function and permits the selection of such an ORF for further development as a screening target. Briefly, these approaches include: generation of temperature sensitive mutations (Weller, S.K., et al., Virology 130:290-305 (1983)), site specific insertion or deletion of a viral gene; a method based on selection of recombinant molecules generated by double recombination through homologous sequencees between intact viral DNA molecules and a DNA fragment containing an insertion or deletion and a selectable marker (Post, L.E., et al., Cell 25:227-32 (1981)), and also by insertional mutagenesis using transposons; a method taking advantage of the random insertion of the DNA phage miniMu into target plasmid DNAs (Jenkins, F.J., et al., Proc. Nat'l. Acad. Sci. USA 82:4773-4777 (1985)). Each of these techniques may have advantages or disadvantages depending on the particular application. The skilled artisan would choose the approach that is the most relevant with the particular end use in mind. For example, some genes might be recognised as essential for infection but in reality are only necessary for the initiation of infection and so their products would represent relatively unattractive targets for antivirals developed to cure established and chronic infections.

Use of these technologies when applied to the ORFs of the present invention enables identification of viral proteins expressed during infection, inhibitors of which would have utility in antiviral therapy.

Glossary:

The following explanations are provided to facilitate understanding of certain terms used frequently herein, particularly in the Examples. The explanations are provided as a convenience and are not limitative of the invention.

HSV-2 BINDING MOLECULE, as used herein, refers to molecules or ions which bind or interact specifically with HSV-2 polypeptides or polynucleotides of the present invention, including, for example, enzyme substrates, cell membrane components and classical receptors. Binding between polypeptides of the invention and such molecules, including binding or interaction molecules may be exclusive to polypeptides of the invention, which is preferred, or it may be highly specific for polypeptides of the invention, which is also preferred, or it may be highly specific to a group of proteins that includes polypeptides of the invention, which is preferred, or it may be specific to several groups of proteins at least one of which includes a polypeptide of the invention. Binding molecules also include antibodies and antibody-derived reagents that bind specifically to polypeptides of the invention.

GENETIC ELEMENT generally means a polynucleotide comprising a region that is important to the viral life cycle, a polynucleotide comprising a region that encodes a polypeptide

or a polynucleotide region that regulates replication, transcription or translation or other processes important to expression of the polypeptide in a host cell, or a polynucleotide comprising both a region that encodes a polypeptide and a region operably linked thereto that regulates expression. Genetic elements may be comprised within a vector that replicates as an episomal element; that is, as a molecule physically independent of the host cell genome. They may be comprised within plasmids. Genetic elements also may be comprised within a host cell genome; not in their natural state but, rather, following manipulation such as isolation, cloning and introduction into a host cell in the form of purified DNA or in a vector, among others.

HOST CELL is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

IDENTITY as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM *J. Applied Math., 48:* 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research 12(1):* 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., *J. Molec. Biol. 215:* 403-410 (1990). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, MD 20894; Altschul, S., *et al., J. Mol. Biol. 215*: 403-410 (1990). As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence it is intended that the nucleotide sequence of the tested polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence. In other words, to

obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. Analogously, by a polypeptide having an amino acid sequence having at least, for example; 95% identity to a reference amino acid sequence is intended that the test amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

ISOLATED means altered "by the hand of man" from its natural state; *i.e.*, that, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a naturally occurring polynucleotide or a polypeptide naturally present in a living organism in its natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. For example, with respect to polynucleotides, the term isolated means that it is separated from the genome and cell in which it naturally occurs. As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in whole organisms. Introduced into host cells in culture or in whole organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides and polypeptides may occur in a composition, such as a media formulations, solutions for introduction of polynucleotides or polypeptides, for example, into cells, compositions or

solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

POLYNUCLEOTIDE(S) generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as used herein refers to, among others, single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded, or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term polynucleotide includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including *inter alia*, simple and complex cells. The term polynucleotide(s) embrace short polynucleotides often referred as oligonucleotides.

POLYPEPTIDES, as used herein, includes all polypeptides as described below. The basic structure of polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other in a linear chain by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types. It will be appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the

terminal amino acids, may be modified in a given polypeptide, either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques which are well known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. Among the known modifications which may be present in polypeptides of the present are, to name an illustrative few, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. Such modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance *PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in *POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS*, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., Meth. Enzymol. 182:626-646 (1990) and Rattan et al., *Protein Synthesis: Posttranslational Modifications and Aging*, Ann. N.Y. Acad. Sci. 663: 48-62 (1992). It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translation natural process and by entirely synthetic methods, as well. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in

polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli* or other cells, prior to proteolytic processing, almost invariably will be N-formylmethionine. During post-translational modification of the peptide, a methionine residue at the NH_2-terminus may be deleted. Accordingly, this invention contemplates the use of both the methionine-containing and the methionineless amino terminal variants of the protein of the invention. The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications in large part will be determined by the host cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as, for example, *E. coli*. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as do mammalian cells and, for this reason, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation. Similar considerations apply to other modifications. It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

VARIANT(S) as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of

polynucleotides and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans.

Deposit:

HSV-2, strain SB5 has been deposited at the American Type Culture Collection under accession number VR-2546 on October 31, 1996.

The deposits referred to herein will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by reference and are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited material, and no such license is hereby granted.

Viral Strain and Genome:

The nucleotide sequences disclosed herein can be obtained by synthetic chemical techniques known in the art or can be obtained from HSV-2, strain SB5 by probing a DNA preparation with probes constructed from the particular sequences disclosed herein. Alternatively, oligonucleotides derived from a disclosed sequence can act as PCR primers in a process of PCR-based cloning of the sequence from a viral genomic source. It is recognised that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The present invention relates to novel HSV-2 polypeptides and polynucleotides encoding same, among other things, as described in greater detail below. The invention relates especially to HSV-2 molecules having the nucleotide and amino acid sequences set out in Tables 1-4 and to the HSV-2 nucleotide and amino acid sequences of the DNA isolatable from Deposit No. ATTC VR-2546, which is herein referred to as "the deposited organism" or as the "DNA of the deposited organism." It will be appreciated that the nucleotide and amino acid sequences set out in Tabled 1-4 were obtained by sequencing the DNA of the deposited organism. Hence, the sequence of the deposited clone is controlling as to any discrepancies between it (and the sequence it encodes) and the sequences of the Tables.

The present invention also relates to additional polynucleotide sequences disclosed herein, which are RNAs transcribed from the DNAs disclosed herein but which may or may not be translated into protein. Such polynucleotides are known in HSV-1 and other herpes viruses.

Polynucleotides

In accordance with one aspect of the present invention, there are provided isolated polynucleotides which encode HSV-2 polypeptides having the deduced amino acid sequence of Tables 1-4. It is preferred that these polynucleotides be one of those set forth in Tables 1, 2 or 3. The skilled artisan can readily determine the polynucleotide sequence of such preferred polynucleotides by reference to the ORF start and stop positions set forth in Tables 1-4.

Using the information provided herein, such as the polynucleotide sequence set out in Tables 1-3, a polynucleotide of the present invention encoding HSV-2 polypeptide may be obtained using standard cloning and screening procedures. To obtain the polynucleotide encoding the protein using the DNA sequences given in Tables 1-3 typically a library of clones of chromosomal DNA of HSV-2 strain SB5 in *E. coli* or some other suitable host is probed with a radiolabelled oligonucleotide, preferably a 17mer or longer, derived from a sequence of Tables 1-3. Clones carrying DNA identical to that of the probe can then be distinguished using high stringency washes. By sequencing the individual clones thus identified with sequencing primers designed from the original sequence it is then possible to extend the sequence in both directions to determine the full gene sequence. Conveniently such sequencing is performed using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook, J. in MOLECULAR CLONING, A Laboratory Manual (2nd edition 1989 Cold Spring Harbor Laboratory. see Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70

The DNA sequences set out in Tables 1, 2 and 3 each contain at least one open reading frame encoding a protein having at least about the number of amino acid residues set forth in Table 1-3. The start and stop codons of each open reading frame are the first three and the last three nuclotides of each polynucleotide set forth in Table 1, 2 and 3.

Certain HSV-2 sequences of the invention are structurally related to sequences encoding other proteins of the herpes family, as shown by comparing the sequences of the Tables with that of sequences reported in the literature. Moreover, certain polynucleotides and polypeptides of the invention are structurally related to known. These proteins exhibit greatest homology to the homologue listed in Tables 1, 2, 3 and 4 from among the known proteins.

The invention provides a polynucleotide sequence identical over its entire length to each coding sequence in Tables 1-3. Also provided by the invention is the coding sequence for the mature polypeptide or a fragment thereof, by itself as well as the coding sequence for the mature polypeptide or a fragment in reading frame with other coding sequence, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence. The polynucleotide may also contain non-coding sequences, including for example, but not limited to non-coding 5'

and 3' sequences, such as the transcribed, non-translated sequences, termination signals, ribosome binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and additional coding sequence which encode additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA 86:* 821-824 (1989), or an HA tag (Wilson *et al.*, *Cell 37:* 767 (1984). Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

The invention also includes polynucleotides of the formula:

$$X\text{-}(R_1)_m\text{-}(R_2)\text{-}(R_3)_n\text{-}Y$$

wherein, at the 5' end of the molecule, X is hydrogen, and at the 3' end of the molecule, Y is hydrogen or a metal, R_1 and R_3 is any nucleic acid residue, n and/or m is an integer between 1 and 3000 or zero, and R_2 is a nucleic acid sequence of the invention, particularly a nucleic acid sequence selected from the group set forth in Tables 1, 2 and 3, as well as a ORF sequence selected from the group set forth in Tables 1, 2, 3 and 4 (as indicated by the reading frame numbering). In the polynucleotide formula above R_2 is oriented so that its 5' end residue is at the left, bound to R_1, and its 3' end residue is at the right, bound to R_3. Any stretch of nucleic acid residues denoted by either R group, where n and/or m is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer. In a preferred embodiment n and/or m is an integer between 1 and 1000, or 2000 or 3000.

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a viral polypeptide and more particularly a polypeptide of the HSV-2 having an amino acid sequence set out in Table 1, 2, 3 or 4. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by integrated phage or an insertion sequence or editing) together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode for variants of the polypeptide having the deduced amino acid sequence of Tables 1, 2, 3 and 4. Variants that are fragments of the polynucleotides of the invention may be used to synthesize full-length polynucleotides of the invention.

Polynucleotides of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be

double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The coding sequence which encodes the polypeptide may be identical to the coding sequence of the polynucleotide shown in Tables 1-4. It also may be a polynucleotide with a different sequence, which, as a result of the redundancy (degeneracy) of the genetic code, encodes the polypeptides of Tables 1-4.

Particularly preferred embodiments are polynucleotides encoding polypeptide variants, that have the amino acid sequence of a polypeptide of Tables 1, 2, 3 and/or 4 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, that do not alter the properties and activities of such polynucleotide.

Further preferred embodiments of the invention are polynucleotides that are at least 50%, 60% or 70% identical over their entire length to a polynucleotide encoding a polypeptide having the amino acid sequence set out in Tables 1, 2, 3 or 4, and polynucleotides that are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 80% identical over its entire length to a polynucleotide encoding a polypeptide of the deposited strain and polynucleotides complementary thereto. In this regard, polynucleotides at least 90% identical over their entire length to the same are particularly preferred, and among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the most preferred.

A preferred embodiment is an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of: a polynucleotide having at least a 50% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of Tables 1, 2, 3, or 4 and obtained from a prokaryotic species other than HSV-2; and a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 50% identical to the amino acid sequence of Tables 1, 2, 3 or 4 and obtained from a prokaryotic species other than HSV-2.

Preferred embodiments are polynucleotides that encode polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by the DNA of Tables 1, 2, 3 or 4.

The invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. An example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about 65°C. Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein.

The invention also provides a polynucleotide consisting essentially of a polynucleotide sequence obtainable by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in Tables 1, 2, 3 or 4 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence or a fragment thereof; and isolating said DNA sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers described elsewhere herein.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as discussed above, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding a polypeptide and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to a polynucleotide set forth in Table 1, 2, 3 or 4. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases. Particularly preferred probes will have at least 30 bases and will have 50 bases or less.

For example, the coding region of each gene that comprises or is comprised by a polynucleotide set forth in Table 1, 2, 3 or 4 may be isolated by screening using a DNA sequence provided in Table 1, 2, 3 or 4 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

Polynucleotides of the invention that are oligonucleotides derived from the a polynucleotide or polypeptide sequence set forth in Table 1, 2, 3 or 4 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in virus in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The invention also provides polynucleotides that may encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

The DNA may also comprise a promoter region which functions to direct the transcription of the mRNA encoding the HSV-2 of this invention . Such promoters may be independently useful to direct the transcription of heterologous genes in recombinant expression systems. Polyadenylation and splicing signal sequences are also present in the polynucleotide sequence and may be useful as gene expression signal in heterologous gene expression vectors and constructs.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for disease, particularly human disease, as further discussed herein relating to polynucleotide assays.

The polynucleotides of the invention that are oligonucleotides may also be used as nucleic acid amplification primers, such as PCR primers, in the process herein described to determine whether or not the HSV-2 genes identified herein in whole or in part are present or transcribed in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

In addition to the uses mentioned above for the polynucleotides of this invention, the following applications are also contemplated by this invention. *Inter alia*, the polynucleotides disclosed herein or portions thereof, may be used as probes to discover

mRNA transcripts synthesized during productive and latent HSV-2 infections, for example by Northern blot, nuclease protection, and primer extension experiments. Novel transcripts in turn can lead to the discovery of new HSV-2 proteins not deducible from the genome sequences directly. The sequences, or portions thereof, may be used to discover antisense inhibitors of virus replication and novel therapeutics based on antisense mechanisms. The sequences, or portions thereof, may be used to prepare novel gene therapy vectors. The sequences or portions thereof may be used as a basis for the generation of DNA- or RNA-containing oligonucleotides designed to form a triplex with duplex DNA, for use as analytical tools, diagnostics or therapeutics. Nucleic acid sequences, or portion thereof, can be used to generate cell lines useful for diagnostics or screening. The DNA sequences can be used to predict restriction enzyme sites useful for replacing the gene in the viral genome with a marker gene such as lac z or green flourescent protein. Such a replacement is useful in defining the biological role of the gene in the viral life cycle. These gene knockout experiments are useful to discover genes which are likely to be high quality drug discovery targets (essential genes) or good locations for foreign genes for the purposes of gene therapy (non-essential genes) through an HSV-2 viral vector. Such gene replacements are also useful for discovering virulence factors, for example by comparing the pathogenicity of the modified virus with the unmodified virus or through the ease of identifying a marker gene such as *lacz*.

In addition to the standard A, G, C, T/U representations for nucleic acid bases, the term "N" is also used. "N" means that any of the four DNA or RNA bases may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a base that when taken in combination with adjacent nucleotide positions, when read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

Polypeptides

The present invention further relates to HSV-2 polypeptides that have the deduced amino acid sequences of the polypeptides defined by amino acid sequence in Tables 1-4.

The invention also relates to fragments, analogs and derivatives of these polypeptides. The terms "fragment," "derivative" and "analog" when referring to the polypeptides of the invention mean a polypeptide which retains essentially the same biological function or activity as such polypeptide. Fragments, derivatives and analogs that retain at least 90% of the biological activity of the native HSV-2 protein are preferred. Fragments, derivatives and analogs that retain at least 95% of the activity of the native HSV-2 protein are preferred. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide. In certain preferred embodiments it is a recombinant polypeptide.

The fragment, derivative or analog of the polypeptides of the invention may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be obtained by those of ordinary skill in the art, from the teachings herein.

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

Further particularly preferred in this regard are variants, analogs, derivatives and fragments having the amino acid sequence of one or more of the HSV-2 polypeptides of the invention, in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the HSV-

2 protein. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polypeptides having the amino acid sequences of Tables 1-4 without substitutions.

The invention also includes polypeptides of the formula:

$$X\text{-}(R_1)_n\text{-}(R_2)\text{-}(R_3)_m\text{-}Y$$

wherein, at the amino terminus, X is hydrogen, and at the carboxyl terminus, Y is hydrogen or a metal, R_1 and R_3 are any amino acid residue, n and/or m is an integer between 1 and 2000 or zero, and R_2 is an amino acid sequence of the invention, particularly an amino acid sequence selected from the group set forth in Tables 1, 2, 3 and 4. In the formula above R_2 is oriented so that its amino terminal residue is at the left, bound to R_1, and its carboxy terminal residue is at the right, bound to R_3. Any stretch of amino acid residues denoted by either R group, where n and/or m is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer. In preferred embodiments n and/or m is an integer between 1 and 1000 or 2000.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polypeptides of the present invention include the polypeptides of Tables 1-4, in particular the mature polypeptide as well as polypeptides which have at least 60%, 70% or 80% identity to one or more of the polypeptides of Tables 1-4 and preferably at least 90% similarity to one or more of the polypeptides of Tables 1-4 and more preferably at least 95% similarity; and still more preferably at least 95% identity to one or more of the polypeptides of Tables 1-4 and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 contiguous amino acids and more preferably at least 50 contiguous amino acids.

In addition to the uses mentioned above for the polypeptides of this invention, the following applications are also contemplated by this invention. *Inter alia*, the polypeptides disclosed herein or portions thereof which have enzymatic activity or structural functionality are useful as a source of those proteins for screening and or therapy. Such polypeptides may be identified by homology for example, to HSV1 polypeptides that code for proteins with known function (e.g., helicases, kinases, proteases). Use of polypeptides of the invention for screening or therapy based upon functionality predicted by homology match is a particularly preferred aspect of this invention. Also the polypeptides derived from the deposited strain ATCC VR-2546 herein can be used for comparison with sequences from other HSV-2 strains in the public domain, for example, comparison of the polypeptides of the invention with strain HG52 may be useful in the discovery of virulence factors, since HG52 is avirulent in mouse and guinea pig infection models and HSV-2 SB5 is virulent. Similarly, public domain homolog from strain MS may be useful in the discovery of

virulence factors since there are major differences in the CNS pathogenesis in animal models between strains MS and SB5.

In addition to the standard single and triple letter representations for amino acids, the term "X" or "Xaa" is also used. "X" and "Xaa" mean that any of the twenty naturally occuring amino acids may appear at such a designated position in the polypeptide sequence.

Fragments

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

Also among preferred embodiments of this aspect of the present invention are polypeptides comprising fragments of HSV-2, most particularly fragments of HSV-2 having the amino acid sequences set out in Tables 1-4, and variants and derivatives thereof.

In this regard, a fragment is a polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned HSV-2 polypeptides and variants or derivatives thereof.

Such fragments may be "free-standing," *i.e.*, not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the presently discussed fragments most preferably form a single continuous region. However, several fragments may be comprised within a single larger polypeptide. For instance, certain preferred embodiments relate to a fragment of a HSV-2 polypeptide of the present comprised within a precursor polypeptide designed for expression in a host and having heterologous pre and pro-polypeptide regions fused to the amino terminus of the HSV-2 fragment and an additional region fused to the carboxyl terminus of the fragment. Therefore, fragments in one aspect of the meaning intended herein, refers to the portion or portions of a fusion polypeptide or fusion protein derived from HSV-2.

Representative examples of polypeptide fragments of the invention, include, for example, those which have from about 5-15, 10-20, 15-40, 30-55, 41-75, 41-80, 41-90, 50-100, 75-100, 90-115, 100-125, and 110-140, 120-150, 200-300, 1-175, 1-600 or 1-1000 amino acids long. Particular examples of polypeptide fragments of the inventions that may be mentioned include fragments of 20-200 amino acids.

In this context about includes the particularly recited range and ranges larger or smaller by several, a few, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes.

Among especially preferred fragments of the invention are truncation mutants of HSV-2. Truncation mutants include HSV-2 polypeptides having the amino acid sequences of Tables 1-4, or of variants or derivatives thereof, except for deletion of a continuous series of residues (that is, a continuous region, part or portion) that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or, as in double truncation mutants, deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Fragments having the size ranges set out above also are preferred embodiments of truncation fragments, which are especially preferred among fragments generally. Degradation forms of the polypeptides of the invention in a host cell are also preferred.

Also preferred in this aspect of the invention are fragments characterized by structural or functional attributes of HSV-2. Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions and high antigenic index regions of HSV-2.

Further preferred regions are those that mediate activities of HSV-2. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of the particular HSV-2 protein , including those with a similar activity or an improved activity, or with a decreased undesirable activity. Routinely one generates the fragment by well-known methods then compares the activity of the fragment to the native protein in a convenient assay such as listed hereinbelow. Highly preferred in this regard are fragments that contain regions that are homologs in sequence, or in position, or in both sequence and to active regions of related polypeptides, such as the related polypeptides set out in Table 1. Among particularly preferred fragments in these regards are truncation mutants, as discussed above. Further preferred polynucleotide fragments are those that are antigenic or immunogenic in an animal, especially in a human.

It will be appreciated that the invention also relates to, among others, polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent conditions, and polynucleotides, such as PCR primers, for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspond to the preferred fragments, as discussed above.

Vectors, host cells, expression:

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate polynucleotides and express polypeptides of the present invention. Introduction of a polynucleotides into the host cell can be affected by calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Polynucelotide constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single or double-stranded phage vector, a single or double-stranded RNA or DNA viral vector. Plasmids generally are designated herein by a lower case p preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art. Starting plasmids disclosed herein are either commercially available, publicly available, or can be constructed from available plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art.

Preferred among vectors, in certain respects, are those for expression of polynucleotides and polypeptides of the present invention. Generally, such vectors comprise *cis*-acting control regions effective for expression in a host operatively linked to the polynucleotide to be expressed. Appropriate *trans*-acting factors either are supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific. Particularly preferred among inducible vectors are vectors that can be induced for expression by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art.

A great variety of expression vectors can be used to express a polypeptide of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, *e.g.*, vectors derived from viral plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Representatives of such promoters include, but are not limited to, the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs.

In general, expression constructs will contain sites for transcription initiation and termination, and, in some instances, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such regions

will operate by controlling transcription, such as transcription factors, repressor binding sites and termination, among others.

Vectors for propagation and expression generally will include selectable markers and amplification regions, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, *E. coli*, streptomyces and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

The following vectors, which are commercially available, are provided by way of example. Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia, and pBR322 (ATCC 37017). Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated that any other plasmid or vector suitable for, for example, introduction, maintenance, propagation or expression of a polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention.

Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("CAT") transcription unit, downstream of restriction site or sites for introducing a candidate promoter fragment; *i.e.*, a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the *cat* gene engenders production of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available, such as pKK232-8 and pCM7. Promoters for expression of polynucleotides of the present invention include not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene.

Among known prokaryotic promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* lacI and lacZ and

promoters, the T3 and T7 promoters, the *gpt* promoter, the lambda PR, PL promoters and the *trp* promoter.

Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Recombinant expression vectors will include, for example, origins of replication, a promoter preferably derived from a highly-expressed or regulatable gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector.

Polynucleotides of the invention, encoding the heterologous structural sequence of a polypeptide of the invention generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to the AUG that initiates translation of the polypeptide to be expressed. Where applicable, a ribosome binding site may be located between the transcription start site and the initiating AUG. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site, where applicable or the 5' end of the transcript and the initiation codon. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal in constructs for use in eukaryotic hosts. Transcription termination signal appropriately disposed at the 3' end of the transcribed region may also be included in the polynucleotide construct.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-or C-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, a region may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability or to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from

immunoglobulin that is useful to solubilize or purify polypeptides. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another protein or part thereof. In drug discovery, for example, proteins have been fused with antibody Fc portions for the purpose of high-throughput screening assays to identify antagonists. See, D. Bennett et al., Journal of Molecular Recognition, 8: 52-58 (1995) and K. Johanson et al., The Journal of Biological Chemistry, 270,(16): 9459-9471 (1995).

Cells typically then are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art.

Mammalian expression vectors may comprise an origin of replication, a suitable promoter and enhancer, and also any necessary polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences that are necessary for expression. In certain preferred embodiments in this regard DNA sequences derived from the SV40 splice sites, and the SV40 polyadenylation sites are used for required non-transcribed genetic elements of these types.

HSV-2 polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, viral, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

HSV-2 polynucleotides and polypeptides may be used in accordance with the present invention for a variety of applications, particularly those that make use of the chemical and

biological properties of HSV-2. Additional applications relate to diagnosis and to treatment of disorders of cells, tissues and organisms. These aspects of the invention are illustrated further by the following discussion.

Polynucleotide assays:

This invention is also related to the use of the HSV-2 polynucleotides to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of HSV-2 polynucleotides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method that can add to, define or allow a diagnosis of a disease. Eukaryotes (herein also "individual(s)"), particularly mammals, and especially humans, infected by HSV-2 may be detected at the DNA or RNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from an individual's cells, tissues, and fluids, such as brain, bone, blood, muscle, cartilage, skin, saliva, urine, semen, and mucous. Tissue biopsy and autopsy material is also preferred for samples from an individual to use in a diagnostic assay. The viral DNA may be used directly for detection or may be amplified enzymatically by using PCR prior to analysis (Saiki et al., Nature 324: 163-166 (1986)). RNA or cDNA may also be used in the same ways. As an example, PCR primers complementary to the nucleic acid encoding HSV-2 can be used to identify and analyze HSV-2 presence and expression. Using PCR, characterization of the strain of virus present in a eukaryote, particularly a mammal, and especially a human, may be made by an analysis of the genotype of the viral gene. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the genotype of a reference sequence. Point mutations can be identified by hybridizing amplified DNA to radiolabeled HSV-2 RNA or alternatively, radiolabeled HSV-2 antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

Sequence differences between a reference gene and genes having mutations also may be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or another amplification method. For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic typing of various strains of virus based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions can be visualized by high

resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, *e.g.*, Myers et al., Science, 230: 1242 (1985)).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (*e.g.*, Cotton et al., Proc. Nat'l. Acad. Sci., USA, 85: 4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (*e.g.*, restriction fragment length polymorphisms ("RFLP") and Southern blotting of genomic DNA.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations also can be detected by *in situ* analysis.

Cells carrying mutations or polymorphisms in the gene of the present invention may also be detected at the DNA level by a variety of techniques, to allow for serotyping, for example. Nucleic acids for diagnosis may be obtained from an infected individual's cells, including but not limited to blood, urine, saliva, tissue biopsy and autopsy material or from virus isolated and cultured from the above or other sources. The viral DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki *et al.*, Nature, 324:163-166 (1986)) prior to analysis. RT-PCR can also be used to detect mutations. It is particularly preferred to used RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA or cDNA may also be used for the same purpose, PCR or RT-PCR. As an example, PCR primers complementary to the nucleic acid encoding HSV-2 can be used to identify and analyze mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled RNA or alternatively, radiolabeled antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures. The primers may be used to amplify the gene isolated from the individual such that the gene may then be subject to various techniques for elucidation of the DNA sequence. In this way, mutations in the DNA sequence may be detected.

Polypeptide assays:

The present invention also relates to diagnostic assays such as quantitative and diagnostic assays for detecting levels of HSV-2 protein in cells and tissues, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in

accordance with the invention for detecting expression of HSV-2 protein compared to normal control tissue samples may be used to detect the presence of an infection. Assay techniques that can be used to determine levels of a protein, such as an HSV-2 protein of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these ELISAs frequently are preferred. An ELISA assay initially comprises preparing an antibody specific to HSV-2, preferably a monoclonal antibody. In addition a reporter antibody generally is prepared which binds to the monoclonal antibody. The reporter antibody is attached to a detectable reagent such as radioactive, fluorescent or enzymatic reagent, in this example horseradish peroxidase enzyme.

Antibodies:

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. The present invention includes, for examples monoclonal and polyclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique known in the art which provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., Nature 256: 495-497 (1975); Kozbor et al., Immunology Today 4: 72 (1983); Cole et al., pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

Alternatively phage display technology could be utilized to select antibody genes with binding activities towards the polypeptide either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-Fbp or from naive libraries

(McCafferty, J. et al., , Nature 348, 552-554 (1990); Marks, J. et al., Biotechnology 10: 779-783 (1992). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. et al., Nature 352, 624-628 (1991).

If two antigen binding domains are present each domain may be directed against a different epitope - termed 'bispecific' antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or purify the polypeptide of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography.

Thus among others, antibodies against HSV-2 may be employed to inhibit and/or treat infections, particularly viral infections, and especially HSV-2 infections as well as to monitor the effectiveness of antibiotic treatment.

Polypeptide derivatives include antigenically, epitopically or immunologically equivalent derivatives which form a particular aspect of this invention. The term "antigenically equivalent derivative" as used herein encompasses a polypeptide or its equivalent which will be specifically recognized by certain antibodies which, when raised to the protein or polypeptide according to the present invention, interfere with the immediate physical interaction between pathogen and mammalian host. The term "immunologically equivalent derivative" as used herein encompasses a peptide or its equivalent which when used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to interfere with the immediate physical interaction between pathogen and mammalian host.

The polypeptide, such as an antigenically or immunologically equivalent derivative or a fusion protein thereof, is used as an antigen to immunize a mouse or other animal such as a rabbit, rat or chicken. The fusion protein may provide stability to the polypeptide. The antigen may be associated, for example by conjugation , with an immunogenic carrier protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively a multiple antigenic peptide comprising multiple copies of the protein or polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

Preferably the antibody or derivative thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanised" ; where the complimentarity determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody , for example as described in Jones, P. et al. Nature 321: 522-525 (1986)or Tempest et al., Biotechnology 9: 266-273 (1991). The above antibody reagents will also be useful for

assessing the biological role of the gene through antibody inhibition studies, immunoprecipitation studies, super-shift experiments and similar techniques. These studies may lead to discovery of novel protein:protein interactions which may be useful drug targets. The above antibody reagents may lead to the identification of novel viral proteins not predicted by the DNA sequence, which in turn may be novel drug targets.

HSV-2 binding molecules and assays:

This invention also provides a method for identification of molecules, such as binding molecules, that bind HSV-2. Genes encoding proteins that bind HSV-2, such as binding proteins, can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance, Coligan et al., *Current Protocols in Immunology 1(2):* Chapter 5 (1991).

For instance, expression cloning may be employed for this purpose. To this end polyadenylated RNA is prepared from a cell expressing HSV-2, a cDNA library is created from this RNA, the library is divided into pools and the pools are transfected individually into cells that are not expressing HSV-2. The transfected cells then are exposed to labeled HSV-2. HSV-2 can be labeled by a variety of well-known techniques including standard methods of radio-iodination or inclusion of a recognition site for a site-specific protein kinase. Following exposure, the cells are fixed and binding of HSV-2 is determined. These procedures conveniently are carried out on glass slides.

Alternatively a labeled ligand can be photoaffinity linked to a cell extract, such as a membrane or a membrane extract, prepared from cells that express a molecule that it binds, such as a binding molecule. Cross-linked material is resolved by polyacrylamide gel electrophoresis ("PAGE") and exposed to X-ray film. The labeled complex containing the ligand-binding can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing can be used to design unique or degenerate oligonucleotide probes to screen cDNA libraries to identify genes encoding the putative binding molecule.

Polypeptides of the invention also can be used to assess HSV-2 binding capacity of HSV-2 binding molecules in cells or in cell-free preparations.

Polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics.

This invention also provides a method of screening drugs to identify those which interfere with the proteins selected as targets herein, which method comprises measuring the

interference of the activity of th protein by a test drug. For example if the protein selected has a catalytic activity, after suitable purification and formulation the activity of the enzyme can be followed by its ability to convert its natural substrates. By incorporating different chemically synthesised test compounds or natural products into such an assay of enzymatic activity one is able to detect those additives which compete with the natural substrate or otherwise inhibit enzymatic activity.

The invention also relates to activators and inhibitors identified thereby.

Another aspect of the invention relates to use of a polynucleotide in genetic immunization, and will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al.*, Hum. Mol. Genet. 1:363 (1992); Manthorpe *et al.*, Hum. Gene Ther. 4:419 (1963)), delivery of DNA complexed with specific protein carriers (Wu *et al.*, J. Biol. Chem. 264:16985 (1989)), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, Proc. Nat'l Acad. Sci. USA, 83:9551 (1986)), encapsulation of DNA in various forms of liposomes (Kaneda *et al.*, Science 243:375 (1989)), particle bombardment (Tang *et al.*, Nature 356:152 (1992)); Eisenbraun *et al.*, DNA Cell Biol. 12:791 (1993)) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, Proc. Nat'l. Acad. Sci. USA 81:5849 (1984)). Suitable promoters for muscle transfection include CMV, RSV, SRa, actin, MCK, alpha globin, adenovirus and dihydrofolate reductase.

In therapy or as a prophylactic, the active agent i.e., the polypeptide, polynucleotide or inhibitor of the invention, may be administered to a patient as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Vaccines:

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal which comprises inoculating the individual with HSV-2 polypeptide, or an antigenic fragment or variant thereof, adequate to produce antibody to protect said individual from infection, particularly HSV-2 infection. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises, through gene therapy, delivering a gene encoding HSV-2, or an antigenic fragment or a variant thereof, for expressing HSV-2, or a fragment or a variant thereof *in vivo* in order to induce an immunological response to produce antibody to protect said individual from disease.

A further aspect of the invention relates to an immunological composition which, when introduced into a host capable or having induced within it an immunological response, induces an immunological response in such host to HSV-2 or a protein coded

therefrom, wherein the composition comprises a recombinant HSV-2 or protein coded therefrom comprising DNA which codes for and expresses an antigen of said HSV-2 or protein coded therefrom.

The HSV-2 or a fragment thereof may be fused with a co-protein which may not by itself produce antibodies, but is capable of stabilizing the first protein and producing a fused protein which will have immunogenic and protective properties. This fused recombinant protein, preferably further comprises an antigenic co-protein, such as Glutathione-S-transferase (GST) or beta-galactosidase, relatively large co-proteins which solubilise the protein and facilitate production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The co-protein may be attached to either the amino or carboxy terminus of the first protein.

The present invention also includes a vaccine formulation which comprises the immunogenic recombinant protein together with a suitable carrier. Since the protein may be broken down in the stomach, it is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation instonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Whilst the invention has been described with reference to acertain HSV-2 polypeptide, it is to be understood that this covers fragments of the naturally occurring protein and similar proteins (for example, having sequence homologies of 75% or greater) with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant protein.

Compositions:

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or the inhibitors. Thus, the polypeptides of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells,

tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

Kits:

The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

Administration:

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. It will be appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions and the like.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

The composition of the invention may be administered by injection to achieve a systemic effect against relevant virus shortly before insertion of an in-dwelling device. Treatment may be continued after surgery during the in-body time of the device. In addition, the composition could also be used to broaden perioperative cover for any surgical technique to prevent viral reactivation.

Alternatively, the composition of the invention may be used to bathe an indwelling device immediately before insertion.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response.

A suitable unit dose for vaccination is 0.5-5 microgram/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks.

With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

In order to facilitate understanding of the following example certain frequently occurring methods and/or terms will be described.

Example 1

Preparation of ultra-purified Herpes simplex 2 virus DNA:

This protocol describes the preparation of herpes simplex virus type 2 strain SB5 DNA for sequencing. It is the combination of two protocols, both of which have been modified. Part one describes the crude isolation of the viral DNA from host cell DNA (Hirt, B., J. Mol. Biol. 26: 365-369. (1967),), and part two describes the ultra-purification of the viral DNA through a cesium chloride (CsCl) gradient (Vinograd J,et al., Proc. Nat'l. Acad. Sci.(USA) 2:902-910(1963)).

I. Separation of viral DNA from host DNA (modified from Hirt[1])

Confluent monolayers of Vero cells (ATCC CCL 81) previously seeded into roller bottles (1×10^8 cells/bottle), were infected with HSV-2 strain SB5 at an MOI = 0.01 in

HBSS. After one hour, the virus innoculum was removed and normal media was added (DMEM, 10% FCS).

Approximately 40-48 hours post- infection, infected monolayers were harvested by scraping, and placed in 10ml of cold 1x PBS. For subsequent steps, three roller bottles of infected cells were combined (3 x 10^8 cells) The cells were spun at 2000g x 5 minutes. The supernatant was removed and to the cell pellet, 25ml of DNA extraction buffer was added (0.25% Triton X-100, 10mM EDTA, 10mM Tris pH 8.0).

The lysate was mixed at room temperature for 10 minutes. Them to the lysate, 1ml of 5M NaCl (0.2M final concentration) was added and allowed to mix another 15 minutes.

The lysate was centrifuged at 10,000g for 30 minutes at 4°C. The supernatant, which contains the viral DNA, was saved and the pellet, which contains mostly chromosomal DNA, was discarded.

To the supernatant, SDS was added to 0.5% final conc. and Proteinase K to 150ug/ml final conc. This was incubated 2 hours at 45°C.

After two hours, 2.5 volumes of 100% ethanol were added. Viral DNA was precipitated overnight at -20°C.

The precipitate was centrifuged at 10,000g for 30 minutes at 4°C. The pellet was washed once with 70% ethanol and air dried for 30 minutes. Then the pellet was resuspended in 250ul of TE (10mM Tris, pH 7.5, 2mM EDTA).

RNase A was added to a final concentration of 10ug/ml and incubated at 37°C for one hour.

SDS and Proteinase K were then added (as above) and incubated overnight at 37°C.

The DNA was phenol extracted 2x, chloroform extracted 1x, and 1/10 volume 3M sodium acetate and 2.5 volumes of 100% ethanol to precipitate were added and allowed to precipitate overnight at -20°C. The next day, The precipitate was spun down at 15,000g x 20 minutes. The pellet was washed 1x with 70% ethanol, briefly air dried and resuspended in 1ml of TE.

II. Ultrapurification of the viral DNA through a CsCl gradient (modified from Vinograd, et al. supra)

A cesium chloride solution of 57% w/w with the prepared DNA from above was made as follows:

To the 1ml of viral DNA prepared above, 9ml of TE was added for a total of exactly 10ml. To this, 13.26g of CsCl was added and dissolved. This solution was added to ultracentrufuge tubes and spun in a VTi 40 rotor at 35,000 rpm for 72 hours at 25°C.

After centrifugation, the tube was mounted on a gradient collector and through a hole pierced in the bottom, 15 drop fractions were collected.

The refractive index of every fourth tube was determined on a refractometer. The viral DNA lies between refractive indicies = 1.403-1.401. Density range for HSV DNA from Goldin A.L, et al., J. Virol. ___: 50-58. Boyant density (ρ) = a $\eta^{25°}$ - b , where coefficients a and b are 10.8601 and 13.4974 respectively for CsCl, η = refractive index. (Isco tables, a handbook of data for biological and physical scientist, Isco, Inc. Lincoln, NE, ninth ed. 1987).

The appropriate fractions were pooled and dialyzed against 3L of TE with frequent changing overnight.

The final DNA prep was concentrated by precipitating with 1/10 volume 3M sodium acetate and 2.5 volumes of 100% ethanol. The DNA was resuspended in TE and the $OD_{260/280}$ reading taken.

The DNA was then subjected to sequencing as provided in Sambrook, J et al. (1989) Chapter 13, supra; or by automated DNA sequencing as per manufacturer's protocols, e.g., Applied Biosystems/Perkin Elmer, Foster City, CA.

Certain preferred individual polynucleotide and polypeptide sequences of the invention are summarized in the following Tables. Tables 1, 2 and 3 represent three different sequencing efforts. Table 4 represents polypeptides encoded by ORFs from Table 3.

Table 1 provides polynucleotides of the invention and polypeptides encoded by ORFs, wherein the polynucleotide start and end position for each ORF is indicated by sequence numbers which correlate to the the polynucleotide sequence referred to above each given polypeptide in the Table. Additionally, each ORF-encoded polypeptide sequence is labeled with the Contig number matching the Contig number of the polynucleotide sequence from which it was encoded. For ORF sequences wherein the start polynucleotide number is larger than the end polynucleotide number, translation of that polypeptide initiates on the nucleotide strand which is complemetary to the strand depicted in the Table. In many cases there is more than one ORF mapped to an individual Contig. Contig assembly was performed using the publicly-available Phrap program, P.Green, University of Washington, WA., U.S.A. ORF prediction was accomplised using the publicly-available GenMark program, Georgia Tech Research Corp., Georgia Tech, GA, U.S.A. Homologies of the polypeptide sequences to known proteins are also indicted. These homologies were

determined using the public database Mpsrch_pp, release 2.1 by J.Collins, Biocomputing Research Unit, University of Edinburgh (distributed by IntelliGenetics, Inc.).

Table 2, obtained from a separately-performed sequencing, provides polynucleotides of the invention and polypeptides encoded by ORFs, wherein the polynucleotide start and end position for ORFs are indicated by sequence numbers which correlate to the polynucleotide sequence referred to above each given polypeptide in the Table. Each ORF-encoded polypeptide sequence is labeled with a Contig number matching the Contig number of the polynucleotide sequence referred to above it, from which it was encoded. For ORF sequences wherein the nucleotide start number is larger than the end number, translation of that polypeptide initiates on the nucleotide strand which is complementary to the strand depicted in the Table. Contig assembly was accomplished using the publicly-available Sequencher 3.0, Gene Codes Corp., Ann Arbor MI, USA, software program. ORF prediction was done using the publicly-available GenMark program (see Table 1). Homologies of the polypeptide sequences to known proteins are indicated. These homologies were determined using the publicly-available Mpsrch program (see Table 1).

Table 3 obtained from a separately-performed sequencing, provides polynucleotides of the invention and polypeptides encoded by ORFs, wherein the polynucleotide start and end positions for each ORF is indicated by sequence numbers which correlate to the polynucleotide sequence referred to above that polypeptide in the Table. Each ORF-encoded polypeptide sequence is labeled with a Contig number matching the Contig number of the polynucleotide sequence appearing above it from which it was encoded. For ORF sequences wherein the start polynucleotide number is larger than the end number, translation of that polypeptide initiates on the nucleotide strand which is complementary to the strand depicted in the Table. Contig assembly was performed using the publicly-available Phrap program, (see Table 1). ORF prediction was accomplished using the publicly-available GenMark software program (see Table 1). Homologies of the polypeptide sequences to known proteins are indicated. These homologies were determined by comparison with public database Mpsrch_pp (see Table 1).

Table 4 provides ORF sequences of polypeptides encoded by the polynucleotide sequences of Table 3 which were predicted by the GenMark program (see Table 1) as having more than a single start site (N-terminal methionyl residue). The Contig numbers and polynucleotide start and end sites for these ORFs correlate to the Contig numbers and polynucleotide sequence numbers of Table 3.

What is claimed is:

1. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

(a) a polynucleotide having at least a 70% identity to a polynucleotide encoding a polypeptide comprising an amino acid sequence of Table 1, 2, 3 or 4;

(b) a polynucleotide having at least a 70% identity to a polynucleotide encoding a mature polypeptide expressed by the gene contained in the HSV-2 of deposited strain VR-2546 that was sequenced to obtain a polynucleotide sequence of Table 1, 2 or 3;

(c) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 70% identical to an amino acid sequence of Table 1, 2, 3 or 4;

(d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and

(e) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a), (b), (c) or (d).

2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.

3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.

4. The polynucleotide of Claim 2 comprising the nucleic acid sequence selected from the group consisting of the nucleic acid sequences set forth in Table 1, 2 and 3.

5. The polynucleotide of Claim 2 which encodes a polypeptide comprising an amino acid sequence sequence selected from the group consisting of the amino acid sequences set forth in Table 1, 2, 3 and 4.

6. A vector comprising the polynucleotide of Claim 1.

7. A host cell comprising the vector of Claim 6.

8. A process for producing a polypeptide comprising expressing in the host cell of Claim 7 a polypeptide encoded by said polynucleotide.

9. A process for producing a polypeptide or fragment thereof comprising culturing a host cell of Claim 7 under conditions sufficient for the production of said polypeptide or fragment.

10. A polypeptide comprising an amino acid sequence which is at least 70% identical to an amino acid sequence selected from the group consisting of the amino acid sequences or fragments thereof set forth in Table 1, 2, 3 and 4.

11. A polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences or fragments thereof set forth in Table 1, 2, 3, and 4.

12. An antibody generated against the polypeptide of claim 10.

What is claimed is:

1. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

(a) a polynucleotide having at least a 70% identity to a polynucleotide encoding a polypeptide comprising an amino acid sequence of Table 1, 2, 3 or 4;

(b) a polynucleotide having at least a 70% identity to a polynucleotide encoding a mature polypeptide expressed by the gene contained in the HSV-2 of deposited strain VR-2546 that was sequenced to obtain a polynucleotide sequence of Table 1, 2 or 3;

(c) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 70% identical to an amino acid sequence of Table 1, 2, 3 or 4;

(d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and

(e) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a), (b), (c) or (d).

2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.

3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.

4. The polynucleotide of Claim 2 comprising the nucleic acid sequence selected from the group consisting of the nucleic acid sequences set forth in Table 1, 2 and 3.

5. The polynucleotide of Claim 2 which encodes a polypeptide comprising an amino acid sequence sequence selected from the group consisting of the amino acid sequences set forth in Table 1, 2, 3 and 4.

6. A vector comprising the polynucleotide of Claim 1.

7. A host cell comprising the vector of Claim 6.

8. A process for producing a polypeptide comprising expressing in the host cell of Claim 7 a polypeptide encoded by said polynucleotide.

9. A process for producing a polypeptide or fragment thereof comprising culturing a host cell of Claim 7 under conditions sufficient for the production of said polypeptide or fragment.

10. A polypeptide comprising an amino acid sequence which is at least 70% identical to an amino acid sequence selected from the group consisting of the amino acid sequences or fragments thereof set forth in Table 1, 2, 3 and 4.

11. A polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences or fragments thereof set forth in Table 1, 2, 3, and 4.

12. An antibody generated against the polypeptide of claim 10.

13. An antagonist or agonist of the activity or expression of the polypeptide of claim 10.

14. A method for the treatment or prevention of disease of an individual comprising administering to the individual a therapeutically effective amount of the polypeptide of claim 10.

15. A method for the treatment of an individual having medical need to inhibit a viral polypeptide comprising administering to the individual a therapeutically effective amount of the antagonist of Claim 13.

16. A process for diagnosing a disease related to expression or activity of the polypeptide of claim 10 in an individual comprising

(a) determining a nucleic acid sequence encoding said polypeptide, and/or

(b) analyzing for the presence or amount of said polypeptide in a sample derived from the individual.

17. A method for identifying compounds which inhibit or activate the polypeptide of claim 10 comprising

(a) contacting a composition comprising the polypeptide with the compound to be screened under conditions to permit interaction between the compound and the polypeptide to assess the interaction of a compound, such interaction being associated with a second component capable of providing a detectable signal in response to the interaction of the polypeptide with the compound; and

(b) determining whether the compound activates or inhibits polypeptide by detecting the presence or absence of the signal generated from the interaction of the compound with the polypeptide.

18. A method for inducing an immunological response in a mammal which comprises inoculating the mammal with the polypeptide of Claim 10, or a variant thereof, adequate to produce antibody and/or T cell immune response to protect said animal from disease.

19. A method of inducing immunological response in a mammal which comprises delivering a nucleic acid vector to direct expression of a polypeptide of Claim 10, or a variant thereof, for expressing said polypeptide *in vivo* in order to induce an immunological response to produce antibody and/ or T cell immune response to protect said animal from disease.

20. The isolated polynucleotide of claim 1 wherein said nucleotide is selected from the group consisting of:

(a) a polynucleotide having at least a 90% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of Table 1, 2, 3 or 4;

(b) a polynucleotide having at least a 90% identity to a polynucleotide encoding the same mature polypeptide expressed by the gene contained in the HSV-2 of the deposited strain VR-2546that was sequenced to obtain a.polynucleotide sequence of Table 1, 2 or 3;

(c) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 90% identical to the amino acid sequence of Table 1, 2, 3 or 4;

(d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and

(e) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a), (b), (c) or (d).

21. The isolated polynucleotide of Claim 1 selected from the group consisting of

(a) a polynucleotide having at least a 95% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of Table 1, 2, 3 or 4;

(b) a polynucleotide having at least a 95% identity to a polynucleotide encoding the same mature polypeptide expressed by the gene contained in the HSV-2 of the deposited strain VR-2546 that was sequenced to obtain a polynucleotide sequence of Table 1, 2 or 3;

(c) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 95% identical to the amino acid sequence of Table 1, 2, 3 or 4;

(d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and

(e) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a), (b), (c) or (d).

22. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

(a) a polynucleotide having at least a 50% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of Table 1, 2, 3 or 4 and obtained from a prokaryotic species other than HSV-2;

(b) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 50% identical to the amino acid sequence of Table 1, 2, 3 or 4 and obtained from a prokaryotic species other than HSV-2; and

(c) a polynucleotide which is complementary to the polynucleotide of (a) or (b).

23. An isolated polypeptide having one of the amino acid sequences given in Table 1, 2, 3 or 4.

24. An isolated nucleic acid encoding one of the amino acid sequences of Claim 1 and nucleic acid sequences capable of hybridizing therewith under stringent conditions.

25. A recombinant vector comprising the nucleic acid sequences of Claim 24 and host cells transformed or transfected therewith.

26. A method of identifying an antiviral compound comprising contacting candidate compounds with a polypeptide of Claim 10 and selecting those compounds capable of inhibiting the bioactivity of said polypeptide.

27. Antiviral compounds identified by the method of Claim 26.

28. An isolated polypeptide having an amino acid sequence or fragment thereof given in Table 1, 2, 3 or 4.

29. An isolated nucleic acid encoding one of the amino acid sequences of Claim 28 and nucleic acid sequences capable of hybridizing therewith under stringent conditions.

30. A method of identifying an antiviral compound comprising contacting candidate compounds with a polypeptide of Claim 28 and selecting those compounds capable of inhibiting the bioactivity of said polypeptide.